Bulletin Technique n° 17
Institut d'Intelligence Artificielle
Université PARIS VIII

# MÉTHODES DE CALCUL POUR UNE ANALYSE INFORMATIQUE DES STRUCTURES SYNTAXIQUES CACHÉES OU VIRTUELLES DE TEXTES POÉTIQUES – (effet "phrases cachées")

Yves LECERF

# - C H A P I T R E 0 -

## OBJET : CALCULS POUR "L'EFFET PHRASES CACHEES"

Les méthodes de calcul qui vont être décrites ont été développées à propos de l'étude de ce que nous avons appelé par ailleurs "l'effet phrases cachées". Cet "effet" (le mot étant pris ici dans l'acception que lui donnent les physiciens pour désigner un fait expérimental insolite) concerne les fragments que l'on peut obtenir dans un texte en procédant à des découpages complètement aléatoires et quelconques, sans en respecter les unités syntaxiques, et bien entendu sans s'obliger à aller d'un point initial de phrase à un point final de phrase.

Le fait expérimental insolite consiste premièrement dans l'observation qu'un nombre relativement important de segments peuvent avoir une cohérence apparente ainsi que des apparences de signification allant souvent à contresens de celle du texte, en sorte que l'on en vient à associer au texte un ensemble de "segments cachés cohérents" qui lui est propre ; et le fait expérimental insolite consiste en second lieu dans l'observation de la possibilité d'une analyse stylistique de cet ensemble de phrases, ensemble qui peut constituer un indicateur plus sensible, plus accessible à des méthodes d'investigations formelles que le texte lui-même.

On trouvera dans nos deux études "L'effet phrases cachées" (Méthodes d'analyse stylistique comportant la prise en considération de structures syntaxiques cachées ou virtuelles) (dans ce même numéro d'ARTINFO/MUSINFO) et "Des poèmes cachés dans des poèmes" (éléments pour une analyse informatique de structures cachées en littérature poetique) (à paraître dans Poétique) des précisions sur ce qu'est "l'effet phrases cachées" et de nombreux exemples de sa manifestation dans divers textes.

Conceptuellement, l'existence expérimentale des phénomènes que nous venons de désigner sous le nom "effet phrases cachées" et leur possibilité d'exploitation en critique stylistique peut être rattachée à l'observation naïve, pure et simple des textes : on aperçoit des segments, on les découpe, on les compare à ceux d'autres textes.

Mais en pratique, l'idée de découper des textes de façon aléatoire ne vient pas naturellement, à moins que des raisons théoriques n'aient laissé prévoir l'apparition d'un ensemble ayant une existence pertinente, comme résultat d'une telle opération.Pour ce qui concerne notre démarche, c'est l'emploi d'un certain type de calcul syntaxique, avec utilisation d'une algèbre de semi anneau sur un magma, qui l'a suggérée ; car dans une telle formalisation des calculs syntaxiques, on voit apparaître avec un rôle extrêmement important un certain outil mathématique multiforme dénommé "produit contracté" ; ce produit contracté apporte notamment une formalisation directe des propriétés de "compilabilité" d'une phrase,et équivaut à un programme universel de compilation.

Or ce même produit contracté, considéré dans des systèmes de pondération légèrement différents, associe justement à toute phrase du texte, l'ensemble de ses "segments cachés cohérents" : autrement dit, l'emploi d'un formalisme de semi-anneau sur un magma conduit tout droit à des concepts exprimant l'effet "phrases cachées".

C'est ce formalisme que nous proposons de décrire ici.

Les lignes qui vont suivre, et qui concernent la définition et les propriétés élémentaires d'une algèbre de semi-anneau sur un magma, paraîtront peut être un peu techniques.

Ce genre de développements peut en général, lorqu'on utilise des structures mathématiques usuelles telles que des anneaux, des corps, des groupes, des espaces vectoriels, etc. pour des applications linguistiques, être évité en faisant report à la littérature mathématique qui énonce les axiomes et propriétés des dites structures.

Mais justement, il n'existe pas à notre connaissance de publication ayant énuméré les propriétés des algèbres de semi-anneau sur un magma : le sujet est donc en principe entièrement non traité, et la rigueur exige que nous énoncions complètement les axiomes de définition de l'algèbre que nous voulons utiliser.

Ce travail préliminaire est nécessaire pour pouvoir passer à une seconde étape qui sera la définition d'un outil permettant une manipulation algébrique de collections de structures et qui sera le "produit contracté". Mais pour commencer il convient d'en rester au "produit simple", non associatif, dans un semi-anneau que nous allons définir sur un magma : ce sera l'objet d'un chapitre I.

La série des chapitres 2 - 3 - 4 qui fera suite sera tout entière relative à l'outil que nous voulions justement en venir à définir et qui est le "produit contracté".

Dans un chapitre 2 on donnera d'abord (2,1) la définition de ce produit: c'est une fonction qui a les propriétés algébriques d'un produit, et qui cependant va constituer un instrument de manipulation de collections de structures, car le "produit contracté" représente toutes les structures associables à une séquence de "mots" donnée.

Dans un paragraphe (2,2) on montrera que le produit contracté est une fonction récursive et on donnera (2,3) la formule de sa définition par récurrence, on étudiera quelques unes de ces propriétés (2,4) et on évoquera le problème de sa représentation par les programmes (2,5)

Dans un chapitre 3 on montrera que le produit contracté coincide avec le fonction de compilation de toute une famille de langages, incluant notamment les context-free.

Dans un chapitre 4, on examinera l'utilisation du produit contracté pour la prise en compte de "l'effet phrases cachées".

# - C H A P I T R E 1 -

## DEFINITIONS D'UNE ALGEBRE DE SEMI-ANNEAU SUR MAGMA : MAGMOIDE

### Définitions

On dira qu'un ensemble E est doté d'une structure de semi anneau sur un magma si cet ensemble est pourvu de deux lois de composition respectivement dénommées addition et multiplication, satisfaisant aux axiomes suivants :

- a) pour la loi de multiplication E est un magma, doté d'un élément neutre (élément unité I) ainsi que d'un zéro (tous produits par un zéro égaux à zéro)

- b) pour la loi d'addition, E est un semi groupe abélien avec élément neutre 0, qui est le zéro de la multiplication

- c) la multiplication est distributive à gauche et à droite par rapport à l'addition.

Loi externe de multiplication par un scalaire : on peut doter en outre l'algèbre de semi anneau définie ci-dessus d'une loi externe de multiplication par un scalaire, dénommée également "Produit" vérifiant les axiomes de :

- distributivité de ce produit par rapport à l'addition scalaire
- associativité pour les multiplications scalaires
- distributivité de ce produit par rapport à l'addition du semi-anneau
- neutralité pour l'élément unité I.x = x

## 1.2 - Propriétés élémentaires de la multiplication du semi-anneau.

On peut rappeler que cette multiplication définit une structure de magma sur E, c'est à dire :

### 1.2.a- La multiplication n'est pas associative.

**Autrement** dit le produit x(y,z) n'est pas généralement égal (pour tous x,y,et z) au produit (x,y)z : autrement dit encore : $x,y,z \rightarrow x(yz) \neq (xy)z$ cette non associativité est particulièrement adaptée à refléter la non associativité des processus d'identifications syntaxiques, par exemple :

Le premier = x
Vide = y
La coupe = z

(le premier vide) la coupe      le premier (vide la coupe)

x y z ≠ x y z

c'est à dire :

$$(x\ y)\ z \neq x\ (y\ z)$$

### 1.2.b - La multiplication n'est pas commutative

Autrement dit : le produit (x,y) n'est pas nécessairement équivalent à (y,x), autrement dit encore :

$$(x,y) \rightarrow xy \neq yx$$

Cette non commutativité est adaptée à refléter la non commutativité des processus d'identification syntaxique, par exemple :

grand = x
homme = y

grand homme      homme grand

x y ≠ y x

c'est à dire :

$$xy \neq yx$$

1.2.c - La multiplication admet un élément zéro (élément permis à droite et à gauche)

Cet élément est tel que pour tout x

$$0.x = x.0 = 0$$

Cet élément est précisément aussi, comme on le verra plus loin, un zéro (élément neutre) pour l'addition : et cette propriété sera utilisée pour la manipulation d'ensembles de structures syntaxiques, lorsqu'il s'agira d'éliminer certaines d'entre elles.

1.2.d - La multiplication admet un élément neutre

Cet élément est donc I tel que pour tout x

$$I.x = x.I = x$$

Cet élément neutre est unique (à noter que zéro n'est pas élément neutre pour la multiplication).

Un tel élément n'a guère de représentation concrète dans le langage (sauf par exemple le blanc qui sépare les mots, mais on n'en voit pas à première vue l'utilisation).

En fait pourtant, il constituera un instrument très utile de calcul pour permettre de créer des expressions contenant à la fois un produit et ses constituants.

Exemple : pour (x,y) : on ajoute I à chacun des constituants et on fait le produit

$$(x + I) (y + I) = xy + x + y$$

Pour des expressions contenant un nombre de facteurs supérieur à deux, on fait intervenir des procédés un peu plus compliqués mais sensiblement de même nature, qui seront décrits plus loin et l'élément unité y joue bien entendu un rôle.

1.2.e - La multiplication est toujours définie

Cette propriété exprime que pour tout couple x et y, d'éléments de E le produit xy existe : elle est une conséquence implicite de la notion d'opération.

Du point de vue des applications à la syntaxe, cette propriété peut surprendre : mais il faut bien voir que l'on dispose de la latitude de faire égal à zéro le résultat de la multiplication, car

$$xy = 0$$

entre dans la catégorie des opérations définies.

1.3 - Propriétés élémentaires de l'addition du semi-anneau et sa distributivité par rapport à la multiplication

Rappelons que cette addition définit une structure de monoïde sur E, il en résulte que :

1.3.a - L'addition est associative

$$(u + v) + w = u + (v + w)$$

La conséquence étant qu'une somme d'un nombre quelconque de termes se note entre deux parenthèses seulement .

$$\begin{aligned}(u + v + w + x + y) &= ((((u + v) + w) + x) + y) \\ &= (u + (v + (w + (x + y)))) \\ &= \text{etc.}\end{aligned}$$

1.3.b - L'addition est commutative (monoïde abélien)

$$u + v = v + u$$

Il en résulte en utilisant l'associativité que l'on peut écrire dans un ordre quelconque les termes d'une somme.

$$u + v + w + x + y = y + x + w + v + u$$
$$= y + u + x + w + v$$
$$= \text{etc.}$$

1.3.c - L'addition admet un élément zéro (élément neutre)

$$u + 0 = 0 + u = u \quad \text{pour tout } u \quad E$$

On démontre facilement que l'élément neutre est unique : donc zéro est le seul élément neutre pour l'addition, à ce propos il y a lieu de rappeler que cet élément zéro (neutre) pour l'addition coïncide avec le zéro (élément permis à gauche et à droite) pour la multiplication, c'est à dire :

$$0\,x = x\,0 = 0$$

et que la notion d'élément permis se distingue de celle d'élément neutre.

1.3.d - Distributivité à gauche par rapport à la multiplication

$$u(x + y) = ux + uy$$

1.3.e - Distributivité à droite par rapport à la multiplication

$$(x + y)u = xu + yu$$

1.3.f - Conséquence pour les formalismes habituels du calcul

Bon nombre d'identités remarquables du calcul algébrique sont transposables pour des calculs dans le semi-anneau, sous réserve de bien préciser l'ordre des opérations de produit.

Exemple : $(x + y)(u + v) = xu + xv + yu + yv$
$((x + y)z)(u + v) = (xz)u + (xz)v + (yz)u + (yz)v$

Il faut toutefois noter que l'addition n'admet pas d'opposé, la multiplication pas d'inverse, ce qui ôte la possibilité de certaines formes de calcul faisant intervenir explicitement ou implicitement ces notions.

L'addition,dotée de son élément zéro, va constituer pour notre objectif (calcul sur des ensembles de structures linguistiques) le moyen de manipuler des collections de segments de phrases, le zéro permettant d'éliminer de ces collections les segments que l'on ne souhaite pas y retenir : on fait en sorte que ces segments soient des produits de valeur nulle, ou inclus dans des produits de valeur nulle.

## 1.4 - Propriétés élémentaires relatives à la loi externe de multiplication, par des scalaires d'un anneau A

Soient les scalaires en lettres grecques, et les éléments de E en lettres latines.

1.4.a - Distributivité par rapport aux scalaires

$$(\alpha + \beta)x = \alpha x + \beta x$$

1.4.b - Distributivité par rapport aux éléments

$$\alpha(x + y) = \alpha x + \alpha y$$

1.4.c - Neutralité pour l'élément unité de l'anneau scalaire A

$$1.x = x$$

1.4.d - Multiplication par le zéro de l'anneau scalaire A

$$0.x = 0$$
$$0.1 = 0$$

1.4.e - Associativité relative au produit des scalaires

$$\alpha(\beta x) = (\alpha\beta)x$$

1.4.f - Associativité généralisée des scalaires

$$(\alpha x)(\beta y) = (\alpha\beta)(xy)$$

L'ensemble de ces propriétés va nous permettre d'effectuer des calculs de comptages sur des structures, une collection de structures pouvant être représentée sous forme polynomiale.

1.4.g - Cas particulier

Un cas particulier remarquable est celui ou le corps K se limite à un seul élément : la multiplication scalaire ne joue plus de rôle, tous les scalaires ayant même poids ; mais une conséquence est que l'addition du semi-anneau devient alors idempotente.

$$\forall x \quad E \qquad x + x = x$$

Ce type particulier d'algèbre est utile pour faire des comptages de formes de structures, lorsque l'on n'a plus besoin de tenir compte du nombre de fois où une même forme de structure est représentée.

1.5 - Magmoïde

On dira qu'un ensemble E est doté d'une structure de Magmoïde avec opération extérieure de multiplication par des pseudo-scalaires d'un ensemble ε si

- E est doté d'une structure de semi-anneau sur un magma

- ε également

- il existe une opération extérieure de multiplication scalaire des éléments E par des éléments scalaires pris dans ε ; opération satisfaisant aux axiomes suivants : distributivité par rapport aux éléments de E ; distributivité par rapport aux éléments aussi de ε ; associativité dans un produit de deux facteurs, des scalaires entre eux et des éléments de E entre eux, chacun des deux produits restant non associatif en lui-même ; neutralité pour l'élément unité de ε, produit nul pour l'élément zéro de ε ou pour celui de E.

# - CHAPITRE 2 -

## DEFINITION DU PRODUIT CONTRACTE SUR LE MAGMA E

### 2.1 - Produit contracté de degré n

La loi de multiplication n'étant pas associative, des expressions telles que :

$$a(bc) \quad \text{et} \quad (ab)c$$

ne sont pas équivalentes. Nous désignerons du nom de $\Pi_3$ (a,b,c) (produit contracté de degré 3) la somme

$$\Pi_3(a,b,c) = a(bc) + (ab)c$$

l'indice 3 sert à rappeler qu'il y a trois variables dans la définition de $\Pi_3$.

De même des expressions telles que :

$$(ab)(cd) \quad \text{et} \quad (a(bc))d \quad \text{et} \quad a((bc)d) \quad \text{et} \quad ((ab)c)d \quad \text{et} \quad a(b(cd))$$

ne sont généralement pas égales et nous désignerons du nom de $\Pi_4(a,b,c,d)$ (produit contracté du degré 4 de a,b,c,d) la somme

$$\Pi_4(a,b,c,d) = (ab)(cd) + (a(bc))d + a((bc)d) + ((ab)c)d + a((bc)d)$$

où l'indice 4 vient rappeler que $\Pi_4$ est définie comme une fonction de 4 variables, dans l'ordre précité.

D'une manière générale, nous désignerons du nom de $\Pi_n(a_1,a_2, \ldots ,a_n)$ la somme de tous les produits possibles faisant intervenir $a_1, a_2, \ldots ,a_n$ dans cet ordre mais avec des parenthetisations différentes.

Pour que la fonction $\Pi_n$ ait un sens avec n = 1 ou avec n = 2, nous poserons

$\Pi_1(a) = a$ quel que soit a
$\Pi_2(a,b) = ab$ quels que soient a et b

en observant du reste que ces définitions conventionnelles ne contreviennent pas à la définition générale donnée pour $\Pi_n$.

## 2.2 - Définition récursive des fonctions $\Pi_n$

### 2.2.a - Définition préliminaire

Opérateurs $(\overset{i}{\rightarrow}\ a_1a_2 \ldots a_n)$ et $(a_1a_2 \ldots a_n\ \overset{j}{\leftarrow})$. On désignera par $(\overset{i}{\rightarrow}\ a_1a_2 \ldots a_n)$ la chaîne composée des i premiers éléments de la suite $a_1a_2 \ldots a_n$. L'opérateur $\overset{i}{\rightarrow}$ n'a de sens que si $i \leq n$ ; il définit une application du monoïde $E^*$ sur lui même. De même on désignera par $(a_1a_2 \ldots a_n\ \overset{j}{\leftarrow})$ la chaîne composée des j derniers éléments de la suite $a_1a_2 \ldots a_n$ ; l'opérateur $\overset{j}{\leftarrow}$ n'a de sens que si $j \leq n$ ; il définit une application du monoïde $E^*$ sur lui même.

Les opérateurs $(\overset{i}{\rightarrow}\ a_1a_2 \ldots a_n)$ et $(a_1a_2 \ldots a_n\ \overset{j}{\leftarrow})$ sont récursifs, et aisément représentables dans les langages de programmation usuels.

### 2.2.b - Définition de $\Pi_n$ à partir des $\Pi_k$ de degré inférieur

La formule de la définition est la suivante :

$$\Pi_n(a_1a_2 \ldots a_n) = \sum_{i=1}^{i=n-1} \Pi_i(\overset{i}{\rightarrow}\ a_1a_2 \ldots a_n)\ \Pi_{n-i}(a_1a_2 \ldots a_n \overset{n-i}{\leftarrow})$$

### 2.2.c - Vérification de l'équivalence des définitions pour les degrés 2,3,4

Degré 2

éléments de calcul : $\Pi_1(a) = a$ et $\Pi_1(b) = b$

La formule donnée ci-dessus pour $\Pi_n$ avec n quelconque devient :

$$\Pi_2(ab) = \Pi_1(a)\ \Pi_1(b)$$

et le résultat $\Pi_2(ab) = ab$ est bien celui déjà donné pour $\Pi_2$

Degré 3

La formule donnée plus haut pour $\Pi_n$ avec n quelconque s'applique avec sommation pour i = 1 et i = 2, soit

$$\Pi_3(abc) = \Pi_1(a)M_2(bc) + \Pi_2(ab)M_1(c)$$

ce qui est bien la définition déjà donnée, car cette somme vaut bien (ab)c + a(bc)

Degré 4

La formule donnée plus haut pour $\Pi_n$ avec n quelconque

$$\Pi_n(a_1a_2 \ldots a_n) = \sum_{i=1}^{i=n-1} \Pi_i(\overset{i}{\rightarrow} a_1a_2 \ldots a_n)\, \Pi_{n-i}(a_1a_2 \ldots a_n \overset{n-i}{\leftarrow})$$

s'applique par sommation pour i = 1, i = 2 et i = 3, soit :

$$\Pi_4(abcd) = \Pi_1(a)M_3(bcd) + \Pi_2(ab)M_2(cd) + \Pi_3(abc)M_1(d)$$

c'est à dire

$$_4(abcd) = a(b(cd) + (bc)d) + (ab)(cd) + (a(bc) + (ab)c)d$$
$$= a(b(cd)) + a((bc)d) + (ab)(cd) + (a(bc))d + ((ab)c)d$$

ce qui est bien la définition déjà donnée.

2.2.d - Equivalence des deux définitions pour n quelconque

Montrons que $\Pi_n(a_1a_2 \ldots a_n)$ défini par la formule de récurrence se compose de tous les produits possibles, faisant intervenir $a_1a_2 \ldots a_n$ dans cet ordre avec des parenthétisations différentes.

- a) Tous les termes de $\Pi_n$ en écriture développée sont des produits faisant intervenir $a_1a_2 \ldots a_n$ dans cet ordre. (En effet si la propriété est vraie pour les degrés inférieurs à n, la structure de la formule fait qu'elle est vraie pour $\Pi_n$)

- b) Ces termes sont tous différents. (En effet si la propriété est vraie pour les degrés inférieurs, la structure de la formule fait qu'elle est vraie pour n)

- c) Tous les termes que l'on peut former à partir de $a_1a_2 \ldots a_n$ dans cet ordre, avec n'importe quelle parenthétisation, sont représentés. (En effet : soit un terme quelconque $(\text{Prod}_n\ a_1a_2 \ldots a_n)$ avec une certaine parenthétisation ou le mot "$\text{prod}_n$" signifie qu'une certaine parenthétisation interne existe, mais que l'on ne l'a pas précisée graphiquement ; et supposons la propriété vraie pour tous les $\Pi_k$ de

degrés $k \leq n$ ; alors le terme quelconque choisi comporte une dernière parenthétisation qui le décompose en produit de deux facteurs $(Prod_i\ a_1 a_2 \ldots a_n)$ et $(Prod_{n-i}\ a_1 a_2 \ldots a_n)$ mais justement $\Pi_i (a_1 a_2 \ldots a_n)$ contient $(Prod_i\ a_1 a_2 \ldots a_n)$ puisque degré $i < n$ et $(\Pi_{n-i} a_1 a_2 \ldots a_n)$ contient $(Prod_{n-i} a_1 a_2 \ldots a_n)$ puisque degré $n-i < n$ donc $\Pi_n$ qui contient $\Pi_i \Pi_{n-i}$ contient $(Prod_i a_1 a_2 \ldots a_n)(Prod_{n-i} a_1 a_2 \ldots a_n)$ donc contient le terme donné $(Prod_n a_1 a_2 \ldots a_n)$

2.3 - Produit contracté, définition générale

Jusqu'à présent nous avons utilisé non pas une fonction produit contracté mais une famille de fonctions $\Pi_n$ produits contractés de degré n (d'ordre n)

Considérons la fonction universelle $\Pi(n, a_1 a_2 \ldots a_n)$ de tous les $\Pi_n$ : la mention de la variable n y joue visiblement un rôle redondant, puisqu'il suffit de compter les autres variables. Donc $\Pi(a_1 a_2 \ldots a_n)$ constituera une notation générale du produit contracté, équivalent à l'ensemble des notations particulières où le degré est mentionné, et la définition s'écrit :

$$\Pi(a_1 a_2 \ldots a_n) = \sum_{i=1}^{i=n-1} \Pi(\overset{i}{\rightarrow}\ a_1 a_2 \ldots a_n)\ \Pi(a_1 a_2 \ldots a_n \overset{n-i}{\leftarrow})$$

2.4 - Propriété de distributivité de produit contracté par rapport à l'addition

2.4.a - Somme sur un terme

$$\Pi(a_1 a_2 \ldots (a_i' + a_i'' + a_i''') \ldots a_n) = \Pi(a_1 a_2 \ldots a_i' \ldots a_n) + \Pi(a_1 a_2 \ldots a_i'' \ldots a_n) + \Pi(a_1 a_2 \ldots a_i''' \ldots a_n)$$

en effet, chacun des deux membres peut être montré équivalent à la somme de tous les produits faisant intervenir les facteurs $a_1 a_2 \ldots a_i \ldots a_n$ dans cet ordre, avec toutes les parenthétisations possibles, la variable $a_i$ prenant les valeurs $a_i'$ ou $a_i''$ ou $a_i'''$ et les valeurs des autres variables étant des valeurs $a_1 a_2 \ldots a_{i-1} a_{i+1} \ldots a_n$ déterminées.

Il y a lieu d'observer que cette règle calque la règle de distributivité d'un produit simple dans un anneau ordinaire où l'on a :

$$a_1 a_2 \ldots (a_i' + a_i'' + a_i''') \ldots a_n = a_1 a_2 \ldots a_i' \ldots a_n + a_1 a_2 \ldots a_i'' \ldots a_n + a_1 a_2 \ldots a_i''' \ldots a_n$$

### 2.4.b - Somme sur plusieurs termes

La règle sera calquée sur celle d'un produit par une somme dans un anneau ordinaire, par exemple :

$$\Pi((a + b)(c + d)) = \Pi ac + \Pi ad + \Pi bc + \Pi bd$$

La procédure générale est la suivante : on considère $\Pi(a_1 a_2 \ldots a_n)$ comme une fonction des variables $(a_1 a_2 \ldots a_n)$ dont chaque facteur peut prendre au choix les valeurs des termes de la somme qui lui correspond ; et l'on fait la somme des $\Pi$ obtenus pour toutes les combinaisons des valeurs de ces variables.

### 2.5 - Programmation de Π, produit contracté

Soit un langage de programmation

- autorisant les définitions "récursives" (appel d'une fonction ou d'une procédure par elle-même)
- comportant des procédures ou fonctions représentatives des calculs "somme" et "produit" du magmoïde
- comportant des procédures ou fonctions représentatives des opérateurs $(\overset{i}{\leftarrow} a_1 a_2 \ldots a_n)$ et $(a_1 a_2 \ldots a_n \overset{i}{\leftarrow})$
- comportant des procédures ou fonctions représentatives de l'opérateur sommation $\sum f(i)g(i)$, pour i allant de $i = 1$ à $i = n-1$
- comportant une procédure ou fonction représentative du "degré" d'une chaîne (nombre de lettres)

Alors la définition

$$\Pi(a_1 a_2 \ldots a_n) = \sum_{i=1}^{i=n-1} \Pi(\overset{i}{\leftarrow} a_1 a_2 \ldots a_n)\, \Pi(a_1 a_2 \ldots a_n \overset{n-i}{\leftarrow})$$

peut se programmer en une seule instruction dans un tel langage (en prenant soin d'exprimer que pour des chaînes de degré 1, $\Pi$ est égal à la chaîne elle-même ; et que pour des chaînes de degré supérieur à 1 $\Pi$ est donné par la formule de sommation).

# - C H A P I T R E 3 -

## OPERATION EXTERIEURE DE PRODUIT PAR UN SCALAIRE OU PSEUDO-SCALAIRE

## LE PRODUIT CONTRACTE COMME FONCTION UNIVERSELLE DE COMPILATION

### 3.1 - Généralités : scalaires, pseudo-scalaires

L'opération extérieure de multiplication par un scalaire ordinaire (élément d'un corps K ou d'un anneau A) s'introduit pour les structures de semi-anneau sur un magma, qui font l'objet de notre étude, de manière analogue à la façon dont on introduit pour la théorie des vecteurs la multiplication extérieure par un scalaire.

Mais il est évident que disposant de la notion de semi-anneau sur un magma, nous l'exploitons beaucoup plus naturellement et surtout beaucoup plus complètement en définissant l'opérateur extérieur comme la multiplication par un "pseudo-scalaire", lequel sera un élément non pas d'un corps, non pas d'un anneau, mais lui aussi d'un semi-anneau sur magma.

Le scalaire ordinaire n'est alors qu'un cas particulier du pseudo-scalaire (c'est le cas où le semi-anneau sur magma devient anneau)

Pour certains calculs, on pourra utiliser successivement une opération extérieure à pseudo-scalaire, puis une opération extérieure à scalaire normal .

### 3.2 - Propriétés élémentaires relatives à la loi extérieure de multiplication

Soient en lettres grecques les pseudo-scalaires pris dans $\varepsilon$ et soient en lettres latines les éléments du semi-anneau de base, E.

Les propriétés sont les suivantes :

- a) distributivité par rapport aux pseudo-scalaires $(\alpha + \beta)x = \alpha x + \beta x$

- b) distributivité par rapport aux éléments de E

$$\alpha(x + y) = \alpha x + \alpha y$$

- c) Neutralité pour l'élément unité de ε

$$1.x = x$$

- d) annulation pour multiplication par l'élément zéro de ε ou de E

$0.x = 0$ et aussi $\alpha.0 = 0$ et aussi $0.0 = 0$

- e) associativité, dans un produit de deux facteurs, des pseudo-scalaires entre eux et des éléments de E entre eux

$$(\alpha x)(\beta y) = (\alpha\beta)(xy)$$

- f) associativité relative à un produit de scalaires

$$\alpha(\beta x) = (\alpha\beta)x$$

- g) quant à la neutralité relative à l'élément unité de E, (qu'impliquerait $\alpha 1 = \alpha$) nous ne l'introduisons pas comme un axiome nécessaire. Cette réserve est la seule qui introduise une dissymétrie entre les rôles de E et de ε.

## 3.3 - Conséquence pour le calcul d'un produit simple et d'un produit contracté

### 3.3.a - Produit simple

Le calcul d'un produit simple du type par exemple

$$\alpha x((\beta y\ \gamma z)\delta u)$$

autorise un regroupement des pseudo-scalaires entre eux et des éléments de E entre eux, les parenthétisations étant les mêmes de part et d'autre

$$\alpha x((\beta y\ \gamma z)\delta u) = \alpha((\beta\gamma)\delta).x((yz)u)$$

Il n'y a pas associativité du produit dans E, ni dans ε, mais on peut regrouper les termes comme ci-dessus ; et inversement, si scalaire et élément de E ont même structure, on peut distribuer le scalaire selon les éléments de E correspondants. Exemple

$$\alpha((\beta.1)\delta).x((yz)u = \alpha x((\beta y.1z)\delta u$$

### 3.3.b - Conséquences pour le calcul d'un produit contracté

Le produit contracté $\Pi\ (\alpha_1 a_1, \alpha_2 a_2\ \ldots\ \alpha_n a_n)$ est égal à la somme de tous les monômes produits que l'on peut former avec $a_1 a_2\ \ldots\ a_n$ dans cet

ordre avec toutes les parenthétisations possibles, chacun d'eux étant précédé d'un produit de $\alpha_1\alpha_2 \ldots \alpha_n$, avec la même parenthétisation.

### 3.4 - Un semi anneau sur magma particulier : cas d'un ensemble ε fini

Si l'ensemble ε est fini, les semi-anneaux sur magma que l'on peut définir sur ε sont isomorphes à des grammaires binaires *context free* dont les éléments non terminaux seraient des éléments de ε.

Considérons en effet toutes les équations de définition d'une telle grammaire

$$\begin{cases} \ldots\ldots \\ \alpha_i \rightarrow \alpha_j\alpha_k \\ \ldots\ldots \end{cases}$$

Ces équations peuvent être considérées comme définissant la table de multiplication d'un semi-anneau

$$\begin{cases} \ldots\ldots \\ \alpha_j\alpha_k \rightarrow \alpha_i \\ \ldots\ldots \end{cases}$$

Les couples non représentés dans cette table seront réputés avoir un produit nul.

Exemple : langage miroir, défini par $E = \{A, S_1, S\}$ et $S \rightarrow AS_1$ et $S_1 \rightarrow SA$

La table de multiplication correspondante s'écrit :

$AS_1 \rightarrow S$

$SA \rightarrow S_1$

et

$AS \rightarrow 0$

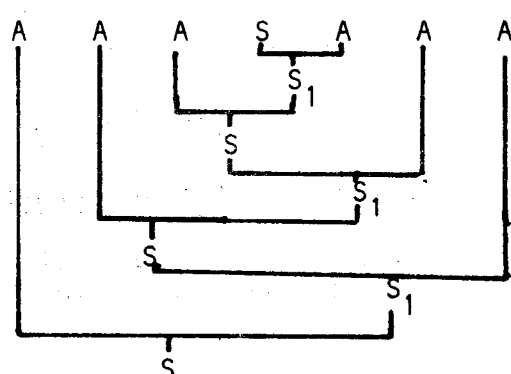
$SS_1 \rightarrow 0$

$S_1A \rightarrow 0$

$S1S \rightarrow 0$

Les seules séquences dont une factorisation produit soit égale à S sont les exemples du type : $A^N S A^N$

Exemple :



On notera que dans $\Pi(A^3SA^3)$ un seul terme est non nul, c'est le terme (A((A((A (S A)) A )) A)) qui correspond au graphe ci-dessus, et dont la valeur est S ; et plus généralement les seuls produits contractés égaux à S seront du type $\Pi(A^NSA^N)$, leur valeur étant fournie par un seul monome (la structure du mot miroir correspondant), et égale à S.

Supposons que nous utilisions ε comme ensemble des pseudo-scalaires pour un ensemble E = {a,c} où a aura comme coéfficient A, et c comme coefficient S. On aura par exemple

$$\Pi((Aa)^3(Sc)(Aa)^3) = S(a((a((ca)a))a))$$

et ce produit nous donne à la fois la structure retenue, et sa valeur grammaticale (Structure en lettres minuscules ; valeur grammaticale par la majuscule en facteur)

## 3.5 - Le produit contracté comme fonction générale de compilation

Le produit contracté apporte une formalisation de la fonction de compilation pour tout langage de structure *context free* binaire : en effet, il suffit d'affecter chaque élément terminal (mot) de la séquence (phrase) à compiler d'un speudo-scalaire égal à la somme des non terminaux susceptibles de lui correspondre ; et d'ajouter un mot supplémentaire, le "point final de la phrase", dont la catégorie sera $P_+$ tel que :

$$SP_+ \rightarrow S_f$$

et

$$XP_+ \rightarrow 0 \qquad \text{pour tout y différent de S}$$

Alors le produit contracté sera égal à S facteur de la somme des structures autorisées par la grammaire.

Il faut signaler que cette propriété de représenter une fonction de compilation s'étend au delà des *context free* ; soit en effet une grammaire infinie du type, par exemple

nombre infini de règles {
..........................................
(Prod $\alpha_i\alpha_j\alpha_k$) (Prod $\alpha_l$ ... $a_r$) → (Prod chaîne fonction récursive du 1er membre)
..........................................

mais dont chaque règle serait une phrase d'un langage dont la grammaire serait par ailleurs finie ; si une telle grammaire régissait les produits pseudo-scalaires, le calcul de compilation resterait possible; or une grammaire comme celle qui vient d'être décrite sort complètement des grammaires *context free*.

# - CHAPITRE 4 -

## LE "PRODUIT CONTRACTE", FORMALISATION DE LA FONCTION "EFFET PHRASES CACHEES"

Nous montrerons la chose uniquement dans le cas de grammaires context free, mais le champ d'application est plus étendu, comme indiqué plus haut.

### 4.1 - L'"effet phrases cachées" limité strictement aux seules phrases cachées de catégorie "S"

Ajoutons à la grammaire context free, qui est supposée donnée dans le cadre de notre exemple, un élément supplémentaire que nous dénommerons "point initial", un élément supplémentaire que nous dnommerons "point final", un élément supplémentaire X ; ajoutons également les relations suivantes :

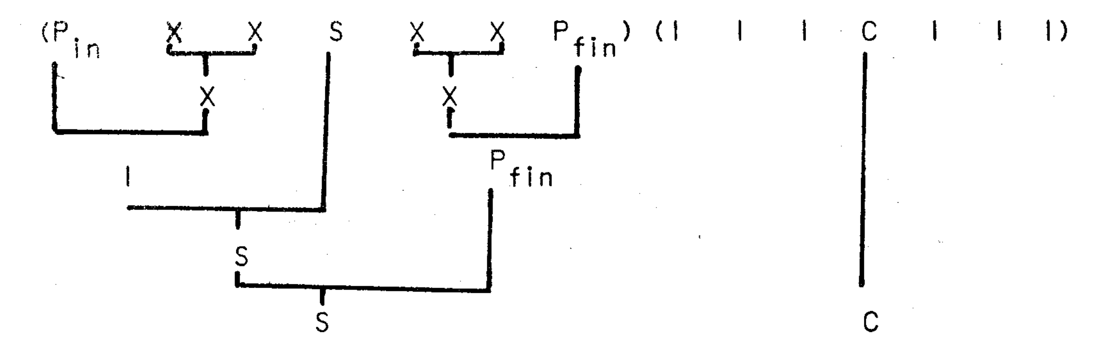
$$XX \rightarrow X$$
$$P_{in}X \rightarrow I \text{ (élément unité)}$$
$$XP_{fin} \rightarrow P_{fin}$$

Associons d'autre part, à tous les mots des phrases à analyser, le terme supplémentaire XI

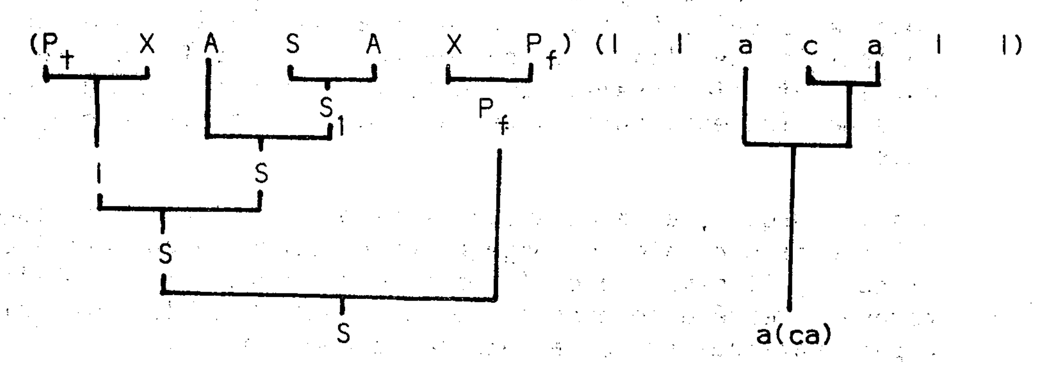
exemple :

$\Pi(a^2ca^2)$ devient ainsi $M(P_{in}(Pa + XI)^2(Sc + XI)(Aa + XI)^2 P_{fin})$

les monomes qui subsistent seront les suivants : d'une part des monomes Sc



ou bien des monomes S(a(ca))




ou bien des monomes $a^2ca^2$ (schéma déjà donné)

Autrement dit, le produit contracté va nous donner (un certain nombre de fois d'ailleurs, mais il importe peu)

S(a((a(ca))a) + a(ca) + c)

qui est la somme des phrases cachées et de la phrase donnée.

## 4.2 - "Effet phrases cachées" étendu à d'autres groupes de mots

On peut souhaiter étendre la notion de"phrases cachées" par exemple à des chaînes nominales.

Les relations mettant en jeu la catégorie X restent les mêmes : mais les relations faisant intervenir le point final sont maintenant au nombre de deux ($P_t$ représentant le point)

$$SP_t \rightarrow S_f \quad \text{et} \quad N_pP_t \rightarrow N_{pf}$$

alors le produit contracté donne, en plus de la structure de phrase complète, toutes les structures connexes partielles (portant sur un segment connexe partiel de phrase) susceptibles d'avoir la valeur $N_p$ ou la valeur S.

## 4.3 - Caractère prévisible de "l'effet phrases cachées" selon le formalisme utilisé

Il faut bien voir que les formules ci-dessus recouvrent une réalité sous jacente plus importante que le "deus es machina" qui permet de conserver tantôt des chaînes phrases, tantôt des chaînes nominales, etc.

Cette réalité sous-jacente est celle des $\Pi_k$ d'ordres $k \leq n$ dont le calcul est nécessaire pour arriver à $\Pi_n$ ; avec cette contrainte que l'on doit pour évaluer $\Pi_n$ calculer non seulement les $\Pi_k$ qui jouent effectivement un rôle, mais encore ceux qui interviennent dans des facteurs nuls, pour vérifier justement que ces facteurs sont nuls.

C'est pourquoi, à la question de savoir si "l'effet phrases cachées" est une face cachée de langage, un discours second caché dérrière le discours, il semble que l'on puisse répondre ceci, que"l'effet phrases cachées" représente dans le langage la face de la compilation, un aspect particulier de la face de la compilation.

-=-=-=-=-=-=-=-=-=-=-=